

Efficient price dynamics in a limit order market: an utility indifference approach*

Masaaki Fukasawa[†]

October 31, 2014

Abstract

We construct an utility-based dynamic asset pricing model for a limit order market. The price is nonlinear in volume and subject to market impact. We solve an optimal hedging problem under the market impact and derive the dynamics of the efficient price, that is, the asset price when a representative liquidity demander follows an optimal strategy. We show that a Pareto efficient allocation is achieved under a completeness condition. We give an explicit representation of the efficient price for several examples. In particular, we observe that the volatility of the asset depends on the convexity of an initial endowment. Further, we observe that an asset price crash is invoked by an endowment shock. We establish a dynamic programming principle under an incomplete framework.

Keywords: utility indifference principle; market impact; efficient price; quadratic backward stochastic differential equation; Burgers' equation.

JEL Classification: G12; G13.

MSC2010 : 91B25; 91G80.

1 Introduction

We are interested in the dynamics of an asset price in a limit order market. We aim at constructing a tractable model which captures endogenously such phenomena observed in actual markets as nonlinearity in liquidation, permanent

*This paper is an output of one of the 2012-2013 Project Researches at Institute of Economic Research, Kyoto University as the Joint Usage and Research Center. The author is grateful to the project members Chiaki Hara, Masaaki Kijima, Katsumasa Nishide, Akihisa Tamura and Yuan Tian for their insightful comments and suggestions. This work was also supported by Japan Society for the Promotion of Science, KAKENHI Grant Number 25245046.

[†]M. Fukasawa

Department of Mathematics, and Center for the Study of Finance and Insurance
Osaka University
1-1 Machikaneyama, Toyonaka, Osaka, 560-0043 JAPAN
Email: fukasawa@math.sci.osaka-u.ac.jp

market impact and a flash crash due to illiquidity. The liquidity issue has been pointed out as a major risk which standard models of financial engineering have not taken into consideration. The failures of several financial institutions are often attributed to the uncovered liquidity risk. The liquidity crisis is a rare event; an exogenous statistical modeling of liquidity costs is therefore not sufficient for preparing ourselves for a future crisis. An economic consideration is required for a deeper understanding of the liquidity risk. This paper provides an utility-based asset pricing model with analytically tractable structure.

In a limit order market, the roles of supplier and demander of liquidity are not symmetric. A liquidity supplier submits a *limit order* that quotes a price for a specified volume of an asset. They can trade with each other by using limit orders to maximize their own utilities. Once an equilibrium is achieved, no more trade occurs among them. However, a liquidity supplier still has an incentive to submit a limit order as long as the corresponding transaction improves her utility. The remained limit orders form a price curve which is a nonlinear function of volume. A liquidity demander submits a *market order* to buy or sell any amount of the asset according to the price curve. Taking a Bertrand-type competition among liquidity suppliers into account, it is then reasonable to begin with modeling the price curve as the utility indifference price for a representative liquidity supplier (RLS hereafter).

If the RLS is risk-neutral, then the utility indifference price of an asset coincides with the expected value of the future cash-flow associated with the asset. The price curve then becomes linear in volume. This simplest framework was adopted by many studies such as Glosten and Milgrom [9]. In this study, we assume the RLS to be risk-averse in managing her inventory, which results in nonlinear pricing with market impact.

This approach differs from the classical works including Garman [8], Amihud and Mendelson [1], Ho and Stoll [11], Ohara and Oldfield [14], where a price quote is a solution of an utility maximization problem for a market maker with exogenously given order-flow. Here, we solve an utility maximization problem for a representative liquidity demander (RLD hereafter). Therefore, an order-flow is endogenously determined. Our model is closely related to the one considered in Bank and Kramkov [2, 3], where they analyzed the market impact of a large trade and formulated a nonlinear stochastic integral as the profit and loss associated with a given strategy of a large trader. Here, we aim at deriving the dynamics of the *efficient price*, that is the asset price when the RLD follows an optimal strategy. Unlike standard optimal investment or hedging problems, the asset price is nonlinear in volume and depends on the trading strategy. If the set of the liquidity demanders consists of price takers and a single large investor, the optimal strategy can be interpreted as the solution of an optimal investment or hedging problem for the large investor under a simplified version of the model introduced by Bank and Kramkov [2, 3]. The model represents permanent market impact, while instantaneous or temporary

market impact models have been extensively considered in the literature. See e.g. Cetin et al. [5], Fukasawa [7], Guéant [10] and the references therein. In the most of the preceding studies, the structure of market impact was exogenously modeled. Here, as in Bank and Kramkov [2, 3], the market impact is endogenous.

We consider a hypothetical market where all traded securities expire at a maturity and yield cash-flow, of which the joint probability distribution is given exogenously. This can be seen as a model of a commodity futures market. A standard theory shows that a commodity future price is determined by an arbitrage relation to the spot price of the commodity. This arbitrage pricing however does not work for a commodity of which the inventory costs are so expensive, such as oil or electricity. A major factor which determines the future price of such a commodity would be the probability distribution of the spot price at the maturity, which is the idea of our model.

In Section 2, we give a rigorous formulation of the problem. In Section 3, we consider a Lévy-driven market as an example which admits explicit results. In particular, we observe that an optimal risk allocation between the RLS and the RLD is achieved in a specific situation. We see also that an endowment shock can invoke a price crash with severe illiquidity. In Section 4, we solve the optimal hedging problem under a condition of completeness, where the optimal risk allocation is achieved. In particular, we construct a perfect hedging strategy under the market impact. We observe that the volatility of the efficient price process depends on the convexity of an initial endowment to be hedged. We see also that a price crash is invoked by a propagated endowment shock. In Section 5, under an incomplete situation, we establish a dynamic programming principle in a discrete time framework and give its applications.

2 Exponential utility indifference pricing

Here we rigorously formulate the problem. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0,1]})$ be a filtered probability space satisfying the usual assumptions with \mathcal{F}_0 composed of the null sets. The time 1 stands for the maturity of all securities and we set $\mathcal{F} = \mathcal{F}_1$. Assume there exists a regular conditional probability measure given \mathcal{F}_t for each $t \in [0, 1)$. By $\mathbb{E}[F|\mathcal{F}_t]$, we always mean the expectation of a random variable F with respect to this regular conditional probability measure. Let

$$S = (S^1, \dots, S^d)^T$$

be an \mathcal{F}_1 -measurable random vector which stands for the cash-flow of d securities at time 1. We assume 0 dividend rate and 0 risk-free rate.

The RLS evaluates future cash-flow F at time 1 by an exponential utility

$$\Pi_t(F) = -\frac{1}{\gamma} \log \mathbb{E}[\exp(-\gamma F)|\mathcal{F}_t] \quad (1)$$

at $t \in [0, 1]$, where $\gamma \geq 0$ is a parameter of risk-aversion. The exponential utility with $\gamma = 0$ is interpreted as the limit of (1) as $\gamma \rightarrow 0$. More specifically, under a suitable condition on F , $\Pi_t(F) = \mathbb{E}[F|\mathcal{F}_t]$ when $\gamma = 0$, which is the case that the RLS is risk-neutral. An important property of the general exponential utility is the cash-invariance:

$$\Pi_t(F + C) = \Pi_t(F) + C$$

for any \mathcal{F}_t -measurable random variable C . Another important property is the quasi-concavity: if $\Pi_t(F^1) \geq 0$ and $\Pi_t(F^2) \geq 0$, then $\Pi_t(\lambda F^1 + (1 - \lambda)F^2) \geq 0$ for any $\lambda \in [0, 1]$. Together with the cash-invariance property, the quasi-concavity implies the concavity:

$$\begin{aligned} & \Pi_t(\lambda F^1 + (1 - \lambda)F^2) - \lambda \Pi_t(F^1) - (1 - \lambda)\Pi_t(F^2) \\ &= \Pi_t(\lambda(F^1 - \Pi_t(F^1)) + (1 - \lambda)(F^2 - \Pi_t(F^2))) \geq 0. \end{aligned}$$

The RLS is initially endowed with a risky asset which yields cash-flow at time 1, denoted by G . If the RLS is holding $z = (z^1, \dots, z^d)$ units of the d securities as her inventory at time $t \in [0, 1]$, then for a volume $y = (y^1, \dots, y^d)$, she quotes a price according to the utility indifference principle:

$$\begin{aligned} P_t(z, y) &= \inf\{p \in \mathbb{R}; \Pi_t(G + zS - yS + p) \geq \Pi_t(G + zS)\} \\ &= \Pi_t(G + zS) - \Pi_t(G + (z - y)S). \end{aligned}$$

Note that the utility indifference price does not depend on the amount of cash held by the RLS due to that her utility is exponential. If the RLS is risk-neutral, then $P_t(z, y) = y\mathbb{E}[S|\mathcal{F}_t]$. Otherwise, the price depends on the inventory z of the securities, which describes permanent market impact. For all t and z , $P_t(z, y)$ is a convex function of y with $P_t(z, 0) = 0$. This implies in particular that

$$-P_t(z, -y) \leq P_t(z, y)$$

for any y and z , which means that the selling price for an amount is higher than or equal to the buying price for the same amount. This represents bid-ask spread that is a measure of market liquidity.

The RLD also evaluates future cash-flow F by an exponential utility

$$U_t(F) = -\frac{1}{c} \log \mathbb{E}[\exp(-cF)|\mathcal{F}_t]$$

at time $t \in [0, 1]$, where $c \geq 0$ is a parameter of risk-aversion. We allow $c = \infty$ by taking the limit $c \rightarrow \infty$. The RLD is initially endowed with a risky asset which yields cash-flow at time 1, denoted by H . Let A be a subset of \mathbb{R}^d with $0 \in A$. We suppose $\Pi_t(G + aS)$ is finite for all $a \in A$ and $t \in [0, 1]$. Let \mathcal{S}_A be the set of the d -dimensional simple predictable processes Y with $Y_0 = 0$ and $Y_t \in A$ for all $t \in [0, 1]$. The RLD is allowed to take any element $Y = (Y^1, \dots, Y^d) \in \mathcal{S}_A$ as her trading strategy, where Y_t^i stands for the number of shares for the i -th security held by the RLD at time t . The price for a volume $y = (y^1, \dots, y^d)$ of

the d securities at time t is $P_t(-Y_t, y)$. This is because the RLS holds $-Y_t$ units of the securities due to the preceding trades with the RLD. Then the profit and loss at time 1 associated with $Y \in \mathcal{S}_A$ is given by

$$I(Y) = Y_1 S - \sum_{0 \leq t < 1} P_t(-Y_t, \Delta Y_t).$$

The problem of the RLD is to maximize the utility of her terminal wealth

$$U_0(H + I(Y))$$

in Y on a suitably extended set of \mathcal{S}_A .

Naturally arise the following questions:

1. Does a dynamic programming principle hold for this utility maximization problem ?
2. How is the optimal strategy Y^* characterized ?
3. How does the *efficient price* $P_t(-Y_t^*, y)$ behave ?
4. How does the *efficient infinitesimal price per unit* (EIPU hereafter)

$$S^* = (S^{*,1}, \dots, S^{*,d}), \quad S_t^{*,i} = \left. \frac{\partial}{\partial y^i} \right|_{y=0} P_t(-Y_t^*, y)$$

behave ?

5. How does the *efficient price convexity*

$$C_t^{ij} = \left. \frac{\partial^2}{\partial y^i \partial y^j} \right|_{y=0} P_t(-Y_t^*, y), \quad i, j = 1, \dots, d$$

behave ?

The EIPU is interpreted as the asset price vector for small investors who are price takers. The efficient price convexity is a measure of infinitesimal illiquidity.

3 Lévy-driven markets

Here we give an example which admits explicit computation. We will see an optimal risk allocation between the RLS and the RLD is achieved in some cases, while not in other cases. Let $d = 1$, $c, \gamma \in (0, \infty)$ and $S = X_1$ for a Lévy process X with $X_0 = 0$. The characteristic function of X_t admits the Lévy-Khintchine representation

$$\frac{1}{t} \log \mathbb{E}[e^{iuX_t}] = iub - \frac{1}{2}\sigma^2 u^2 + \int (e^{iux} - 1 - iuh(x))\nu(dx),$$

where $b, \sigma \in \mathbb{R}$, $h(x) = x1_{\{|x| \leq 1\}}$ and ν is a measure on \mathbb{R} with

$$\int 1 \wedge x^2 \nu(dx) < \infty.$$

Assume that there exists an interval $U \subset \mathbb{R}$ such that

$$\int e^{-ux} 1_{\{|x| > 1\}} \nu(dx) < \infty$$

for all $u \in U$. Then,

$$\int |e^{-ux} - 1 + uh(x)| \nu(dx) < \infty$$

for all $u \in U$ and by Theorem 25.3 of Sato [17],

$$\Pi_t(zX_1) = zX_t + \frac{1-t}{\gamma} \kappa(\gamma z), \quad (2)$$

for all $z \in U/\gamma$, where

$$\kappa(u) = bu - \frac{1}{2} \sigma^2 u^2 - \int (e^{-ux} - 1 + uh(x)) \nu(dx).$$

Since Π_0 is a concave functional, κ is a concave function with $\kappa(0) = 0$. For example,

$$\kappa(u) = bu - \frac{1}{2} \sigma^2 u^2, \quad U = \mathbb{R}$$

if $X_t = bt + \sigma W_t$, where W is a standard Brownian motion. We have

$$\kappa(u) = \beta \log(1 + u/\alpha), \quad U = (-\alpha, \infty)$$

if X is a gamma process with X_1 following the gamma distribution with rate $\alpha > 0$ and shape $\beta > 0$. Another example is

$$\kappa(u) = ru^\alpha, \quad U = [0, \infty)$$

when X is a one-sided stable process with rate $r > 0$ and exponent $\alpha \in (0, 1)$.

Suppose $G = aS$ for $a \in U/\gamma$ and $A \subset a - U/\gamma$. Then, by (2),

$$P_t(z, y) = yX_t + \frac{1-t}{\gamma} (\kappa(\gamma(a+z)) - \kappa(\gamma(a+z-y))),$$

for $z \in -A$ and $y \in A+z$ and so,

$$\begin{aligned} P_t(-Y_t, \Delta Y_t) &= X_t \Delta Y_t - \frac{1-t}{\gamma} \Delta \kappa(\gamma(a - Y_t)) \\ &= X_{t-} \Delta Y_t + \Delta X_t \Delta Y_t - \frac{1-t}{\gamma} \Delta \kappa(\gamma(a - Y_t)) \end{aligned}$$

for $Y \in \mathcal{S}_A$. By integration by parts,

$$Y_1 X_1 - \sum_{0 \leq t < 1} P_t(-Y_t, \Delta Y_t) = \int_0^1 Y_t dX_t + \frac{1}{\gamma} \int_0^1 (\kappa(\gamma(a - Y_t)) - \kappa(\gamma a)) dt. \quad (3)$$

Note that if 0 is an interior point of U , then κ is differentiable at 0 and by letting $\gamma \rightarrow 0$,

$$P_t(z, y) \rightarrow y \tilde{X}_t, \quad \tilde{X}_t = X_t + (1 - t)\mathbb{E}[X_1].$$

The process \tilde{X} is a martingale and it was the asset price process when the RLS was risk-neutral. The right hand side of (3) can be written as

$$\int_0^1 Y_t d\tilde{X}_t + \int_0^1 Y_t \left\{ \kappa'(0) - \frac{\kappa(\gamma a) - \kappa(\gamma(a - Y_t))}{\gamma Y_t} \right\} dt. \quad (4)$$

The first term is of the familiar form of profit and loss associated with the trading strategy Y when \tilde{X} was the price process. The second term can be both positive and negative and converges to 0 as $\gamma \rightarrow 0$.

Now, consider H of the form

$$H = h + \int_0^1 H'_t dX_t,$$

where $h \in \mathbb{R}$ and H' is a simple predictable bounded process independent of X . This specification is just to obtain an explicit result. If one interprets dX_t as an economic signal at time t , then H'_t determines its contribution to the initial endowment H .

If $c(H' + Y)$ takes values in U , then the terminal wealth of the RLD is

$$\begin{aligned} & h + \int_0^1 (H'_t + Y_t) dX_t + \frac{1}{\gamma} \int_0^1 (\kappa(\gamma(a - Y_t)) - \kappa(\gamma a)) dt \\ &= h + \int_0^1 (H'_t + Y_t) dX_t - \frac{1}{c} \int_0^1 \kappa(c(H'_t + Y_t)) dt \\ & \quad + \frac{1}{c} \int_0^1 \kappa(c(H'_t + Y_t)) dt + \frac{1}{\gamma} \int_0^1 (\kappa(\gamma(a - Y_t)) - \kappa(\gamma a)) dt \\ &\leq h + \int_0^1 (H'_t + Y_t) dX_t - \frac{1}{c} \int_0^1 \kappa(c(H'_t + Y_t)) dt \\ & \quad + \frac{c + \gamma}{c\gamma} \int_0^1 \kappa\left(\frac{c\gamma}{c + \gamma}(a + H'_t)\right) dt - \frac{1}{\gamma} \kappa(\gamma a). \end{aligned}$$

The last inequality follows from the convexity of κ and the equality is attained when

$$Y_t = Y_t^* := \frac{\gamma}{c + \gamma} a - \frac{c}{c + \gamma} H'_t. \quad (5)$$

Since H' is independent of X , we have

$$\mathbb{E} \left[\exp \left\{ -c \int_0^1 (H'_t + Y_t) dX_t + \int_0^1 \kappa(c(H'_t + Y_t)) dt \right\} \middle| H' \right] = 1$$

if, say, Y is bounded. See Kallsen and Shiryaev [13]. If A is bounded but enough large for \mathcal{S}_A to include Y^* given by (5), then

$$\begin{aligned} \max_{Y \in \mathcal{S}_A} U_0(H + I(Y)) &= U_0(H + I(Y^*)) \\ &= U_0 \left(h + \frac{c + \gamma}{c\gamma} \int_0^1 \kappa \left(\frac{c\gamma}{c + \gamma} (a + H'_t) \right) dt \right) - \frac{1}{\gamma} \kappa(\gamma a). \end{aligned}$$

To interpret this identity, let us introduce

$$U_{H'}^*(F) = -\frac{c + \gamma}{c\gamma} \log \mathbb{E} \left[\exp \left\{ -\frac{c\gamma}{c + \gamma} F \right\} \middle| H' \right]$$

for random variable F . Then,

$$U_{H'}^*(G + H) = h + \frac{c + \gamma}{c\gamma} \int_0^1 \kappa \left(\frac{c\gamma}{c + \gamma} (a + H'_t) \right) dt$$

and so,

$$\max_{Y \in \mathcal{S}_A} U_0(H + I(Y)) = U_0(U_{H'}^*(G + H)) - \Pi_0(G). \quad (6)$$

Note that

$$\frac{c\gamma}{c + \gamma} < c, \quad U_0(U_{H'}^*(G + H)) \leq U_0^*(U_{H'}^*(G + H)) = U_0^*(G + H),$$

where

$$U_t^*(F) = -\frac{c + \gamma}{c\gamma} \log \mathbb{E} \left[\exp \left\{ -\frac{c\gamma}{c + \gamma} F \right\} \middle| \mathcal{F}_t \right] \quad (7)$$

for random variable F . This functional U^* is known as the aggregated utility under optimal risk allocation (see e.g., Barrieu and El Karoui [4]). The above computation shows that the optimal risk allocation between the RLS and the RLD is achieved by a dynamic trading strategy Y^* if H' is deterministic, while some utility can be lost otherwise.

Since the optimal strategy Y^* of the RLD is given by (5), the efficient price and the EIPU are respectively

$$\begin{aligned} P_t(-Y_t^*, y) &= yX_t + \frac{1-t}{\gamma} \left(\kappa \left(\frac{c\gamma}{c + \gamma} (a + H'_t) \right) - \kappa \left(\frac{c\gamma}{c + \gamma} (a + H'_t) - \gamma y \right) \right), \\ S_t^* &= \frac{\partial}{\partial y} \bigg|_{y=0} P_t(-Y_t^*, y) = X_t + (1-t) \kappa' \left(\frac{c\gamma}{c + \gamma} (a + H'_t) \right) \end{aligned}$$

if κ is differentiable. The latter can be rewritten as

$$S_t^* = \tilde{X}_t - (1-t) \left(\kappa'(0) - \kappa' \left(\frac{c\gamma}{c+\gamma} (a + H_t') \right) \right).$$

Since the first term is a martingale, the second term is interpreted as the *risk premium*. On intervals where H' is constant, we have

$$dS_t^* = \left(\kappa'(0) - \kappa' \left(\frac{c\gamma}{c+\gamma} (a + H_t') \right) \right) dt + d\tilde{X}_t.$$

Since κ is concave, the risk premium is nonnegative if the aggregated initial endowment $a + H_t'$ is nonnegative. A positive (resp. negative) endowment shock $\Delta H'$ induces a nonpositive (resp. nonnegative) jump of S^* and a nonnegative (resp. nonpositive) jump of the risk premium. The efficient price convexity is computed as

$$\frac{\partial^2}{\partial y^2} \Big|_{y=0} P_t(-Y_t^*, y) = -\gamma(1-t)\kappa'' \left(\frac{c\gamma}{c+\gamma} (a + H_t') \right) \geq 0$$

if κ is twice differentiable. This implies that the illiquidity level strongly depends on the convexity of κ and so, the distribution tail of S . If k'' is one more times differentiable and the Lévy measure ν is spectrally negative, then $k''' < 0$ and the skewness of S is negative. In this case, since $-\kappa''$ is increasing, a positive endowment shock $\Delta H' > 0$ induces a positive jump of the infinitesimal illiquidity, in other words, a wider bid-ask spread, in addition to a drop of S^* .

4 Complete markets

4.1 The completeness condition

Here, we assume the filtration $\{\mathcal{F}_t\}$ is generated by a k -dimensional standard Brownian motion W . We solve the problem of the RLD under a completeness condition specified below. We set $A = \mathbb{R}^d$ for brevity. Let \mathcal{L} be the set of random variables F such that for all $a > 0$, $\mathbb{E}[\exp(a|F|)] < \infty$. Note that \mathcal{L} is a linear space. We assume that G, H and S are elements of \mathcal{L} . Let Π^y be a continuous modification of $\Pi(G - yS)$ and

$$Z_t^y = (Z_t^{y,1}, \dots, Z_t^{y,k}), \quad Z_t^{y,i} = -\frac{d}{dt} \langle \Pi^y, W^i \rangle_t$$

for $y \in \mathbb{R}^d$. The key assumption of the section is the existence of a map

$$Z : \Omega \times [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}^k$$

with the following properties:

1. For each $t \in [0, 1]$, the restriction

$$Z|_{[0,t]} : \Omega \times [0, t] \times \mathbb{R}^d \rightarrow \mathbb{R}^k$$

is $\mathcal{F}_t \otimes \mathcal{B}([0, t]) \times \mathcal{B}(\mathbb{R}^d)$ -measurable.

2. There exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that

$$Z(\omega, t, y) = Z_t^y(\omega)$$

for all $(\omega, t, y) \in \Omega_0 \times [0, 1] \times \mathbb{R}^d$.

3. There exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that $Z(\omega, t, y)$ is continuous in y for each $(\omega, t) \in \Omega_0 \times [0, 1]$.

4. There exists a map

$$Y^\dagger : \Omega \times [0, 1] \times \mathbb{R}^k \rightarrow \mathbb{R}^d$$

such that the restriction

$$Y^\dagger|_{[0,t]} : \Omega \times [0, t] \times \mathbb{R}^k \rightarrow \mathbb{R}^d$$

is $\mathcal{F}_t \otimes \mathcal{B}([0, t]) \otimes \mathcal{B}(\mathbb{R}^k)$ -measurable for each $t \in [0, 1]$ and there exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that

$$Z(\omega, t, Y^\dagger(\omega, t, z)) = z$$

for each $(\omega, t, z) \in \Omega_0 \times [0, 1] \times \mathbb{R}^k$.

We call this assumption the *completeness condition* hereafter.

The terminology is justified by considering the risk-neutral case; when $\gamma = 0$ and G and S are square-integrable, then

$$\Pi_t^y = \mathbb{E}[G|\mathcal{F}_t] + y\mathbb{E}[S|\mathcal{F}_t] = \int_0^t (G'_t + yS'_t) dW_t$$

almost surely by the Itô representation theorem, where G' and S' are progressively measurable processes which are \mathbb{R}^k -valued and $\mathbb{R}^d \otimes \mathbb{R}^k$ -valued respectively. In this case, the asset price is linear in volume and

$$dS_t^* = S'_t dW_t.$$

Since

$$Z_t^y = -G' - yS'_t,$$

the completeness condition is satisfied if S'_t has rank k as a $d \times k$ matrix and admits a left inverse matrix which is progressively measurable as a process.

Another elementary example is the case $G = a + bW_1$ and $S = \alpha + \beta W_1$ with $a \in \mathbb{R}, b \in \mathbb{R}^k, \alpha \in \mathbb{R}^d$ and $\beta \in \mathbb{R}^d \otimes \mathbb{R}^k$. Since

$$\Pi_t(G - yS) = a - y\alpha + (b - y\beta)W_t - \frac{\gamma}{2}(1-t)|b - y\beta|^2,$$

the completeness condition is satisfied if and only if β has rank k .

Lemma 1 For $Y \in \mathcal{S}_A$, we have

$$\mathcal{I}(Y) = G - \Pi_0(G) - \int_0^1 \frac{\gamma}{2} |Z_t^Y|^2 dt + \int_0^1 Z_t^Y dW_t,$$

where $Z_t^Y(\omega) = Z(\omega, t, Y_t(\omega))$.

Proof: By the completeness condition, we identify $Z_t^y(\omega) = Z(\omega, t, y)$. Let $Y \in \mathcal{S}_A$. Denote by

$$0 \leq \tau_1 < \tau_2 < \dots$$

the stopping times when Y jumps. We have $\tau_n = 1$ for sufficiently large n almost surely. Remember that $Y_{\tau_{j+1}}$ is \mathcal{F}_{τ_j} -measurable. Note that

$$\begin{aligned} \mathcal{I}(Y) &= G - \Pi_1(G - Y_1 S) - \sum_{0 \leq t < 1} (\Pi_t(G - Y_t S) - \Pi_t(G - Y_{t+} S)) \\ &= G - \Pi_{\tau_1}^0(G) - \sum_{j=1}^{\infty} (\Pi_{\tau_{j+1}}^{(j)} - \Pi_{\tau_j}^{(j)}), \end{aligned}$$

where $\Pi^{(j)} = \Pi^y$ with $y = Y_{\tau_{j+1}}$. As is well-known, (Π^y, Z^y) is the unique solution of the backward stochastic differential equation (BSDE hereafter)

$$G - yS = \Pi_t^y + \int_t^1 \frac{\gamma}{2} |Z_s^y|^2 ds - \int_t^1 Z_s^y dW_s.$$

To see this, it is enough to observe that $M = \exp\{-\gamma \Pi^y\}$ is a local martingale with $M_1 = \exp\{-\gamma(G - yS)\}$ and then, apply the Itô representation theorem. This BSDE representation implies that

$$\Pi_t^y - \Pi_s^y = \int_s^t \frac{\gamma}{2} |Z_u^y|^2 du - \int_s^t Z_u^y dW_u$$

and so, the result follows. ////

By Lemma 1, we can extend the definition of $\mathcal{I}(Y)$ onto the set \mathcal{S} of the progressively measurable processes Y with

$$\int_0^1 |Z_t^Y|^2 dt < \infty$$

almost surely and

$$-\int_0^1 \frac{\gamma}{2} |Z_t^Y|^2 dt + \int_0^1 Z_t^Y dW_t \in \mathcal{L}.$$

The problem is then to find an optimal strategy $Y \in \mathcal{S}$.

4.2 The hedging strategy for the RLD with $c = \infty$

Lemma 2 *For any $H \in \mathcal{L}$, there exists $Y \in \mathcal{S}$ such that*

$$-H = \Pi_0(G) - \Pi_0(G + H) + I(Y).$$

Proof: This is a perfect hedging problem and amounts to finding the solution (V, Z^*) of the BSDE

$$G + H = V_t + \int_t^1 \frac{\gamma}{2} |Z_s^*|^2 ds - \int_t^1 Z_s^* dW_s.$$

As before, this quadratic BSDE admits an explicit solution given by

$$V_t = \Pi_t(G + H), \quad Z^* = (Z^{*,1}, \dots, Z^{*,k}), \quad Z_t^{*,i} = -\frac{d}{dt} \langle V, W^i \rangle_t.$$

By the completeness condition, we can define a progressively measurable process Y^* by

$$Y_t^*(\omega) = Y^\dagger(\omega, t, Z_t^*(\omega)) \quad (8)$$

to have

$$Z^{Y^*} = Z^*.$$

////

By Lemma 2, any claim $-H$ is perfectly replicated by a strategy $Y \in \mathcal{S}$ under the completeness condition. The *replication price* of $-H$ is given by

$$p(-H) := \Pi_0(G) - \Pi_0(G + H). \quad (9)$$

This is convex in $-H$, which can be interpreted as a diversification of the liquidity risk.

Theorem 1 *If $c = \infty$, then*

$$\max_{Y \in \mathcal{S}} U_0(H + I(Y)) = U_0(H + I(Y^*)) = \Pi_0(G + H) - \Pi_0(G),$$

where Y^* is given by (8).

Proof: Recall a well-known representation

$$U_0(F) = \inf_{Q \sim \mathbb{P}} \left\{ \mathbb{E}[F \frac{dQ}{d\mathbb{P}}] + \frac{1}{c} \mathbb{E}[\frac{dQ}{d\mathbb{P}} \log \frac{dQ}{d\mathbb{P}}] \right\}$$

for $c \in (0, \infty)$. Take the limit $c \rightarrow \infty$, to have

$$U_0(F) = \inf_{Q \sim \mathbb{P}} \mathbb{E}[F \frac{dQ}{d\mathbb{P}}] = \text{ess.inf } F$$

for any bounded random variable F (See e.g., Delbaen [6]). This means that the RLD is extremely risk-averse and tries to offset her initial endowment H almost surely. By Lemma 2,

$$H + I(Y^*) = \Pi_0(G + H) - \Pi_0(G)$$

is constant and so,

$$\sup_{Y \in \mathcal{S}} U_0(H + I(Y)) \geq U_0(H + I(Y^*)) = \Pi_0(G + H) - \Pi_0(G).$$

Suppose that there exists $Y \in \mathcal{S}$ such that

$$\text{ess.inf}(H + I(Y)) > \Pi_0(G + H) - \Pi_0(G). \quad (10)$$

Since

$$G - I(Y) = \Pi_0(G) + \int_0^1 \frac{\gamma}{2} |Z_t^Y|^2 dt - \int_0^1 Z_t^Y dW_t,$$

we have

$$\Pi_0(G - I(Y)) = \Pi_0(G) \quad (11)$$

by the uniqueness of the solution of the BSDE. From this and (10), we deduce a contradiction as follows:

$$\Pi_0(G) = \Pi_0(G - I(Y)) < \Pi_0(G + H) - (\Pi_0(G + H) - \Pi_0(G)) = \Pi_0(G).$$

////

4.3 The optimal strategy for the RLD with $c < \infty$

Here we extend Theorem 1 to the case $c < \infty$. First we state a version of the result referred as Borch's theorem in Barrieu and El Karoui [4].

Lemma 3 *If $c < \infty$, then for any $G, H \in \mathcal{L}$,*

$$\sup_{F \in \mathcal{L}} \{U_0(F) + \Pi_0(G + H - F)\} = U_0(F^*) + \Pi_0(G + H - F^*) = U_0^*(G + H),$$

where

$$F^* = \frac{\gamma}{c + \gamma}(G + H)$$

and U^* is defined by (7).

Proof: Since

$$F \mapsto U_0(F) + \Pi_0(G + H - F)$$

is concave, it suffices to observe that

$$\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \{U_0(F^* + \epsilon X) + \Pi_0(G + H - F^* - \epsilon X)\} = 0$$

for all $X \in \mathcal{L}$. ////

Theorem 2 *If $c < \infty$,*

$$\max_{Y \in \mathcal{S}} U_0(H + I(Y)) = U_0(H + I(Y^*)) = U_0^*(G + H) - \Pi_0(G),$$

where Y^ is given by*

$$Y_t^*(\omega) = Y^\dagger(\omega, t, \frac{c}{c+\gamma} Z_t^\dagger(\omega)), \quad Z_t^{\dagger,i} = -\frac{d}{dt} \langle U^\dagger, W^i \rangle_t, \quad i = 1, \dots, d$$

and U^\dagger is the continuous modification of $U^(G + H)$.*

Proof: By definition,

$$H + I(Y^*) = G + H - \Pi_0(G) + \frac{c}{c+\gamma} \left\{ - \int_0^1 \frac{1}{2} \frac{c\gamma}{c+\gamma} |Z_t^\dagger|^2 dt + \int_0^1 Z_t^\dagger dW_t \right\}.$$

Since (U^\dagger, Z^\dagger) is the solution of the BSDE

$$G + H = U_t^\dagger + \int_t^1 \frac{1}{2} \frac{c\gamma}{c+\gamma} |Z_s^\dagger|^2 ds - \int_t^1 Z_s^\dagger dW_s,$$

we have

$$H + I(Y^*) = \frac{\gamma}{c+\gamma} (G + H) + \frac{c}{c+\gamma} U_0^*(G + H) - \Pi_0(G).$$

It follows then,

$$\begin{aligned} U_0(H + I(Y^*)) &= \frac{\gamma}{c+\gamma} U_0^*(G + H) + \frac{c}{c+\gamma} U_0^*(G + H) - \Pi_0(G) \\ &= U_0^*(G + H) - \Pi_0(G). \end{aligned}$$

It remains to show that

$$\sup_{Y \in \mathcal{S}} U_0(H + I(Y)) \leq U_0(H + I(Y^*)).$$

As before, we have (11) for all $Y \in \mathcal{S}$. Therefore,

$$U_0(H + I(Y)) \leq U_0^*(G + H) - \Pi_0(G + H - (H + I(Y))) = U_0^*(G + H) - \Pi_0(G)$$

by Lemma 3. ////

Since $c/(c+\gamma) \rightarrow 1$ and so, $U_0^*(G + H) \rightarrow \Pi_0(G + H)$ as $c \rightarrow \infty$, Theorem 2 extends Theorem 1. As mentioned in the previous section, $U_0^*(G + H)$ is the aggregated utility of the aggregated initial endowment. The above result implies that an optimal risk allocation (Pareto allocation) is achieved by the dynamic trading strategy Y^* for general $G, H \in \mathcal{L}$ under the completeness condition.

4.4 Markov models

Here we assume in addition that

$$S = s(W_1), \quad G = g(W_1), \quad H = h(W_1)$$

with Borel functions $s : \mathbb{R}^k \rightarrow \mathbb{R}^d$ and $g, h : \mathbb{R}^k \rightarrow \mathbb{R}$ to extract a more tractable structure of the optimal strategy. Define functions $v : [0, 1] \times \mathbb{R}^k \rightarrow \mathbb{R}$ and $p : [0, 1] \times \mathbb{R}^k \times \mathbb{R}^d \rightarrow \mathbb{R}$ as

$$\begin{aligned} v(t, w) &= -\frac{c + \gamma}{c\gamma} \log \mathbb{E} \left[\exp \left\{ -\frac{c\gamma}{c + \gamma} (G + H) \right\} \middle| W_t = w \right], \\ p(t, w, y) &= -\frac{1}{\gamma} \log \mathbb{E} \left[\exp \{ -\gamma(G - yS) \} \middle| W_t = w \right]. \end{aligned} \quad (12)$$

By the assumption on the filtration, we have

$$v(t, W_t) = \Pi_t^*(G + H), \quad p(t, W_t, y) = \Pi_t(G - yS). \quad (13)$$

As is well-known and easily checked, v and p are the solutions of the partial differential equations (PDE hereafter)

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{1}{2} \left\{ \Delta v - \frac{c\gamma}{c + \gamma} |\nabla v|^2 \right\} &= 0, \quad v(1, w) = g(w) + h(w), \\ \frac{\partial p}{\partial t} + \frac{1}{2} \left\{ \Delta p - \gamma |\nabla p|^2 \right\} &= 0, \quad p(1, w, y) = g(w) - ys(w), \end{aligned}$$

where

$$\Delta = \sum_{i=1}^k \frac{\partial^2}{\partial w_i^2}, \quad \nabla = \left(\frac{\partial}{\partial w_1}, \dots, \frac{\partial}{\partial w_k} \right).$$

The completeness condition can be restated as follows: there exists a measurable function $y^\dagger : \mathbb{R}^k \times [0, 1] \times \mathbb{R}^k \rightarrow \mathbb{R}^d$ such that

$$-\nabla p(t, w, y^\dagger(w, t, z)) = z \quad (14)$$

for each $(z, t, w) \in \mathbb{R}^k \times [0, 1] \times \mathbb{R}^k$. The optimal strategy given by Theorems 1 and 2 is

$$Y_t^* = y^\dagger(W_t, t, -\frac{c}{c + \gamma} \nabla v(t, W_t)).$$

The efficient price, the efficient price per unit and the efficient price convexity are given by

$$\begin{aligned} P_t(-Y_t^*, y) &= p(t, W_t, Y_t^*) - p(t, W_t, Y_t^* + y), \\ S_t^* &= -\frac{\partial p}{\partial y^i}(t, W_t, Y_t^*), \quad i = 1, \dots, d, \\ C_t^{ij} &= -\frac{\partial^2 p}{\partial y^i \partial y^j}(t, W_t, Y_t^*), \quad i, j = 1, \dots, d. \end{aligned}$$

4.5 The volatility of the EIPU

Here we study through a simple example how the volatility of the EIPU is determined by model parameters. Let $k = d = 1$, $G = gS$, $S = \mu + \sigma W_1$ and

$$H = aW_1 + \frac{b}{2}W_1^2 = a\frac{S - \mu}{\sigma} + \frac{b}{2}\left(\frac{S - \mu}{\sigma}\right)^2$$

for constants $g, \mu, \sigma, h, a, b \in \mathbb{R}$. We assume $c + \gamma + c\gamma b > 0$ and $\sigma \neq 0$. Although H does not belong to \mathcal{L} , it is not difficult to see the preceding results are extended to this specific model. In this case, S follows a normal distribution and

$$G + H = g\mu + (g\sigma + a)W_1 + \frac{b}{2}|W_1|^2.$$

It is straightforward to have

$$\begin{aligned} v(t, w) &= g\mu + (g\sigma + a)w + \frac{b}{2}|w|^2 - \frac{c\gamma}{2} \frac{(g\sigma + a + bw)^2(1-t)}{c + \gamma + c\gamma b(1-t)} \\ &\quad + \frac{c + \gamma}{2c\gamma} \log\left(1 + \frac{c\gamma}{c + \gamma} b(1-t)\right), \\ p(t, w, y) &= (g - y)\mu + (g - y)\sigma w - \frac{\gamma}{2}(g - y)^2\sigma^2(1-t). \end{aligned}$$

It follows then that

$$\frac{\partial p}{\partial w}(t, w, y) = (g - y)\sigma, \quad \frac{\partial v}{\partial w} = (g\sigma + a + bw) \frac{c + \gamma}{c + \gamma + c\gamma b(1-t)}.$$

The completeness condition is therefore satisfied and we have

$$Y_t^* = g - \frac{g\sigma + a + bW_t}{\sigma} \frac{c}{c + \gamma + c\gamma b(1-t)}.$$

The efficient price, the EIPU and the efficient price convexity are given by

$$\begin{aligned} P_t(-Y_t^*, y) &= y(\mu + \sigma W_t) + \frac{\gamma}{2}\sigma^2(1-t) \left\{ (g - Y_t^* - y)^2 - (g - Y_t^*)^2 \right\}, \\ S_t^* &= \mu + \sigma W_t - \gamma\sigma^2(1-t)(g - Y_t^*) = \mu + \sigma \frac{(c + \gamma)W_t - (g\sigma + a)c\gamma(1-t)}{c + \gamma + c\gamma b(1-t)}, \\ C_t &= \gamma\sigma^2(1-t). \end{aligned}$$

The volatility of S^* is therefore given by

$$\frac{d}{dt}\langle S^* \rangle_t = \frac{\sigma^2(c + \gamma)^2}{(c + \gamma + c\gamma b(1-t))^2}.$$

The volatility monotonically depends on b , the convexity of $-H$ as a function of S . In case b is negative, or equivalently, if the RLD has to hedge an European payoff $-H$ which is a convex function of S , then the volatility is larger in case $\gamma > 0$ than in the risk-neutral case $\gamma = 0$. This means that economic signals are amplified in a market where hedgers of convex payoffs are dominant.

4.6 Ride a shock wave

Each of the PDEs in (12) is a KPZ equation and by differentiating in w , we obtain (backward) Navier-Stokes equations for irrotational flow

$$\begin{aligned}\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u &= \frac{1}{2} \frac{c\gamma}{c+\gamma} \nabla |u|^2, \quad u(1, w) = \nabla(g+h)(w), \\ \frac{\partial q}{\partial t} + \frac{1}{2}\Delta q &= \frac{1}{2} \gamma \nabla |q|^2, \quad q(1, w) = \nabla(g-ys)(w)\end{aligned}$$

for $u = \nabla v$ and $q = \nabla p$ when s, g and h are differentiable. In case $k = 1$, they are Burgers equations

$$\begin{aligned}\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial w^2} &= \frac{c\gamma}{c+\gamma} u \frac{\partial u}{\partial w}, \quad u(1, w) = \frac{\partial}{\partial w}(g+h)(w), \\ \frac{\partial q}{\partial t} + \frac{1}{2} \frac{\partial^2 q}{\partial w^2} &= \gamma q \frac{\partial q}{\partial w}, \quad q(1, w) = \frac{\partial}{\partial w}(g-ys)(w).\end{aligned}\tag{15}$$

A nontrivial explicit solution for the Burgers equation is known; it is easy to see

$$u(t, w) = 1 - \tanh(a(w - w_c) - a^2(1 - t)), \quad a = \frac{c\gamma}{c+\gamma},$$

is a solution with

$$(g+h)(w) = w - \frac{1}{a} \log(\cosh(a(w - w_c))) + b$$

where w_c and b can be arbitrary constants. The solution is unique; see Hopf [12]. Remark that if $b = -w_c$, then

$$-(g+h)(w) \rightarrow 2(w_c - w)_+$$

as $a \rightarrow \infty$. Therefore in case $g = 0$ and $S = \mu - \sigma W_1$ with constants $\mu, \sigma > 0$, the problem of RLD is to hedge a payoff which is close to a call option payoff:

$$-H \approx \frac{2}{\sigma} (S - (\mu - \sigma w_c))_+.$$

The completeness condition is satisfied and

$$Y_t^* = \frac{1}{\sigma} \frac{c}{c+\gamma} \frac{\partial v}{\partial w}(t, W_t) = \frac{1}{\sigma} \frac{c}{c+\gamma} (1 - \tanh(a(W_t - w_c) - a^2(1 - t))).$$

It follows then that

$$S_t^* = \mu - \sigma W_t + \sigma(1 - t)a(1 - \tanh(a(W_t - w_c) - a^2(1 - t))).$$

In case that $a = c\gamma/(c+\gamma)$ is large, the function u has a steep slope around $w - a(1 - t) = w_c$. This means that when $W_t - a(1 - t)$ approaches to w_c to cross

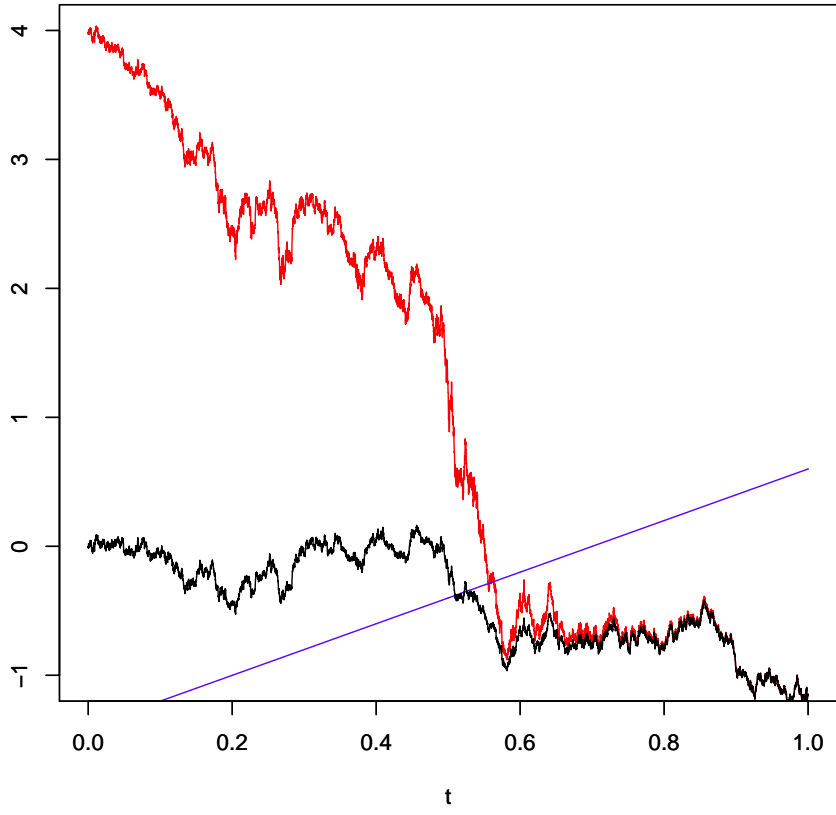


Figure 1: A crash of the EIPU (red line) when the factor process $-W$ (black line) meets a shock wave (blue line for the steepest position $-w_c - a(1 - t)$): $\mu = 0$, $\sigma = 1$, $a = 2$ and $w_c = -0.6$.

it from below, the EIPU S^* exhibits a drastic drop. See Figure 1 for a sample path. This can be interpreted as an asset price crash invoked by a propagated endowment shock.

The Burgers equation models fluid dynamics and is known to form a shock wave. More precisely, an inviscid Burgers equation can form a shock wave with a singularity. A viscous Burgers equation like (15) does not form a singularity but can form a steep shape because as $a \rightarrow \infty$, the equation tends to be inviscid. As in the above explicit case, the EIPU will crash when the factor process W meets a shock wave. The shock wave does not have a singularity but has a steep slope when a is large, or equivalently, when both the RLS and RLD are strongly risk averse. A more detailed analysis on this shock wave model would be worth to be done; it remains for future research.

5 Dynamic Programming principle

5.1 Incomplete markets

Here we show a version of dynamic programming principle holds in our problem. Let $D_n = \{0, 1/n, \dots, (n-1)/n\}$ with an integer $n > 0$ and denote by \mathcal{S}_A^n the set of strategies $Y \in \mathcal{S}_A$ such that $\Delta Y_t \neq 0$ only if $t \in D_n$. In this section, we consider the maximization of $U_0(H + I(Y))$ subject to $Y \in \mathcal{S}_A^n$. As the following lemma shows, this problem coincides with the original one if the filtration $\{\mathcal{F}_t\}$ is discrete in the sense that $\mathcal{F}_t = \mathcal{F}_{[nt]/n}$ for all $t \in [0, 1]$. This means that information is updated only at a discrete time set.

Lemma 4 *If $\mathcal{F}_t = \mathcal{F}_{[nt]/n}$ for all $t \in [0, 1]$, then*

$$\sup_{Y \in \mathcal{S}_A} U_0(H + I(Y)) = \sup_{Y \in \mathcal{S}_A^n} U_0(H + I(Y)).$$

Proof: For $t \in [j/n, (j+1)/n)$, putting $\tau = j/n$,

$$\begin{aligned} & P_\tau(-Y_\tau, \eta) + P_t(-Y_\tau - \eta, y) \\ &= \Pi_\tau(G_m - Y_\tau S) - \Pi_\tau(G_m - (Y_\tau + \eta)) \\ & \quad + \Pi_t(G_m - (Y_\tau + \eta)S) - \Pi_t(G_m - (Y_\tau + \eta + y)S) \\ &= \Pi_\tau(G_m - Y_\tau S) - \Pi_\tau(G_m - (Y_\tau + \eta + y)S) \\ &= P_\tau(-Y_\tau, \eta + y), \end{aligned}$$

which means that buying η units at time τ and y units at time t is equivalent to buying $\eta + y$ units at time τ . ////

Proposition 1 *Define $\{V_t^n\}$ recursively by*

$$V_t^n(x, z) = \begin{cases} H + x - zS & \text{if } t \geq 1, \\ \sup_{y \in A+z} U_{[nt]/n}(V_{([nt]+1)/n}^n(x - P_{[nt]/n}(z, y), -z + y)) & \text{otherwise} \end{cases}$$

for $(x, z) \in \mathbb{R} \times (-A)$. If A is compact, then

$$V_0^n(0, 0) = \max_{Y \in \mathcal{S}_A^n} U_0(H + \mathcal{I}(A))$$

Proof: First note that $V^n(x, z)$ is constant on $[j/n, (j+1)/n)$ for each (x, z) and $j = 0, 1, \dots, n-1$. Since

$$U_0(F) = U_0 \circ U_{1/n} \cdots \circ U_{(n-1)/n}(F)$$

for any random variable F and

$$\begin{aligned} & U_{(j-1)/n}(V_{j/n}^n(-\sum_{0 \leq t < j/n} P_t(-Y_t, \Delta Y_t), Y_{j/n})) \\ & \leq V_{(j-1)/n}^n(-\sum_{0 \leq t < (j-1)/n} P_t(-Y_t, \Delta Y_t), Y_{(j-1)/n}) \end{aligned}$$

for any $Y \in \mathcal{S}_A^n$ and $j = 1, 2, \dots, n$, we obtain

$$V_0^n(0, 0) \geq \sup_{Y \in \mathcal{S}_A^n} U_0(H + Y_1 S - \sum_{0 \leq t < 1} P_t(-Y_t, \Delta Y_t)).$$

On the other hand, since A is compact, there exists a sequence $\{Y_j^*\}$ adapted to $\{\mathcal{F}_{j/n}\}$ such that

$$V_0^n(0, 0) = U_0(H + Y_1 S - \sum_{0 \leq t < 1} P_t(-Y_t, \Delta Y_t))$$

with

$$Y_t = \sum_{j \leq [nt]} Y_j^*$$

by the measurable selection theorem (see Parthasarathy [15]).

////

Lemma 5 Let F be a random variable and $t \in D_n$. Define a sup-convolution operator Ψ_t as

$$\Psi_t(F) = \sup\{U_t(F - L) + \Pi_t(L); L \in \{\Pi_{t+1/n}(G - yS); y \in A\}\}.$$

Then, for any \mathcal{F}_t -measurable random vector $Z = (Z^1, \dots, Z^d)$,

$$\begin{aligned} & \sup_{y \in A+Z} U_t(F - \Pi_{t+1/n}(G + (Z - y)S + P_t(Z, y))) \\ & = \Psi_t(F) - \Pi_t(G + ZS). \end{aligned}$$

Proof: Since

$$\begin{aligned} & \Pi_{t+1/n}(G + (Z - y)S + P_t(Z, y)) \\ & = \Pi_{t+1/n}(G + (Z - y)S) + P_t(Z, y) \\ & = \Pi_{t+1/n}(G + (Z - y)S) + \Pi_t(G + ZS) - \Pi_t(G + (Z - y)S) \end{aligned}$$

and $\Pi_t(G + (Z - y)S) = \Pi_t(\Pi_{t+1/n}(G + (Z - y)S))$,

$$\begin{aligned} & U_t(F - \Pi_{t+1/n}(G + (Z - y)S + P_t(Z, y))) + \Pi_t(G + ZS) \\ &= U_t(F - \Pi_{t+1/n}(G + (Z - y)S)) + \Pi_t(\Pi_{t+1/n}(G + (Z - y)S)). \end{aligned}$$

Taking sup in y , we obtain the representation. ////

Proposition 2

$$V_0^n(0, 0) + \Pi_0(G) = \Psi_0 \circ \Psi_{1/n} \circ \dots \Psi_{(n-1)/n}(G + H).$$

Proof: Note that

$$V_1^n(x, z) = H + x - zS = x + G + H - \Pi_1(G + zS).$$

Then, applying the previous lemma with $F = G + H$,

$$\begin{aligned} V_{(n-1)/n}^n(x, z) &= x + \sup_{y \in A+z} U_{(n-1)/n}(F - \Pi_1(G + (z - y)S + P_{(n-1)/n}(z, y))) \\ &= x + \Psi_{(n-1)/n}(F) - \Pi_{(n-1)/n}(G + zS). \end{aligned}$$

Repeat this to get the representation. ////

5.2 No rebalancing solution

As a direct application of the above results, let us consider a special case that there exists $y^* \in A$ such that

$$G - y^*S = \frac{c}{c + \gamma'}(G + H), \quad H + y^*S = \frac{\gamma}{c + \gamma'}(G + H).$$

This holds, for example, both G and H are proportional to S and A is sufficiently large. Assume G , H and S are bounded and A is compact. Applying the preceding results, we have

$$\max_{Y \in S_A^n} U_0(H + I(Y)) = \Psi_0 \circ \dots \Psi_{(n-1)/n}(G + H) - \Pi_0(G).$$

Let us observe that the supremum is attained by

$$Y_t = y^* \quad t \in (0, 1], \quad Y_0 = 0.$$

This means that the RLD buys y^* units of the securities at time 0 and holds them until time 1. The supremum in

$$\Psi_{(n-1)/n}(G + H) = \sup_{y \in A} \{U_{(n-1)/n}(H + yS) + \Pi_{(n-1)/n}(G - yS)\}$$

is attained by $y^* \in A$ because

$$y \mapsto U_{(n-1)/n}(H + yS) + \Pi_{(n-1)/n}(G - yS)$$

is concave and

$$\frac{\partial}{\partial y^i} \{U_{(n-1)/n}(H + yS) + \Pi_{(n-1)/n}((G - yS))\} \Big|_{y=y^*} = 0$$

for all $i = 1, \dots, d$. Further we have

$$\begin{aligned} \Psi_{(n-1)/n}(G + H) &= -\frac{c + \gamma}{c\gamma} \log \mathbb{E} \left[\exp \left\{ -\frac{c\gamma}{c + \gamma} (G + H) \right\} \middle| \mathcal{F}_{(n-1)/n} \right] \\ &= \frac{c + \gamma}{c} \Pi_{(n-1)/n} \left(\frac{c}{c + \gamma} (G + H) \right). \end{aligned}$$

Then we go to the next step. We have

$$\begin{aligned} &\Psi_{(n-2)/n}(\Psi_{(n-1)/n}(G + H)) \\ &= \sup_{y \in A} \{U_{(n-2)/n}(F - \Pi_{(n-1)/n}(G - yS)) + \Pi_{(n-2)/n}(G - yS)\} \\ &\leq \sup_{H \in \mathcal{L}} \{U_{(n-2)/2}(F - \Pi^* - L) + \Pi_{(n-2)/2}(\Pi^* + L)\}, \end{aligned}$$

where

$$\begin{aligned} F &= \frac{c + \gamma}{c} \Pi_{(n-1)/n} \left(\frac{c}{c + \gamma} (G + H) \right) \\ \Pi^* &= \Pi_{(n-1)/n}(G - y^*S) = \Pi_{(n-1)/n} \left(\frac{c}{c + \gamma} (G + H) \right) = \frac{c}{c + \gamma} F \end{aligned}$$

and \mathcal{L} is the linear space spanned by $\Pi_{(n-1)/n}(G - yS)$, $y \in A$. The function

$$\epsilon \mapsto U_{(n-2)/2}(F - \Pi^* - \epsilon L) + \Pi_{(n-2)/2}(\Pi^* + \epsilon L)$$

is concave for each $L \in \mathcal{L}$ and its derivative in ϵ at $\epsilon = 0$ vanishes due to that $c(F - \Pi^*) = \gamma \Pi^*$. Therefore

$$\begin{aligned} &\sup_{y \in A} \{U_{(n-2)/n}(F - \Pi_{(n-1)/n}(G - yS)) + \Pi_{(n-2)/n}(G - yS)\} \\ &\leq U_{(n-2)/2}(F - \Pi^*) + \Pi_{(n-2)/2}(\Pi^*) \end{aligned}$$

and the upper bound is attained when $y = y^*$. Consequently we get

$$\Psi_{(n-2)/n}(\Psi_{(n-1)/n}(G + H)) = -\frac{c + \gamma}{c\gamma} \log \mathbb{E} \left[\exp \left\{ -\frac{c\gamma}{c + \gamma} (G + H) \right\} \middle| \mathcal{F}_{(n-2)/n} \right]$$

and apparently this argument can be repeated to conclude that the optimal strategy to hold y^* units from the beginning.

Consequently, the optimal strategy is $Y_t^* = y^*$ for $t > 0$, which implies

$$\begin{aligned}\max_{Y \in S_A^y} U_0(H + I(Y)) &= U_0(H + y^*S - P_0(0, y^*)) \\ &= U_0(H + y^*S) + \Pi_0(G - y^*S) - \Pi_0(G) \\ &= U_0^*(G + H) - \Pi_0(G).\end{aligned}$$

This extends (6) and means that the optimal risk allocation is achieved. Denote by Q the probability measure defined by

$$\frac{dQ}{dP} = \frac{\Gamma}{\mathbb{E}[\Gamma]}, \quad \Gamma = \exp \left\{ -\frac{c\gamma}{c + \gamma}(G + H) \right\}.$$

Then,

$$S_t^* = \frac{\mathbb{E}[S \exp\{-\gamma(G - y^*S)\} | \mathcal{F}_t]}{\mathbb{E}[\exp\{-\gamma(G - y^*S)\} | \mathcal{F}_t]} = \mathbb{E}^Q[S | \mathcal{F}_t]. \quad (16)$$

Further, the efficient price convexity coincides with the conditional covariance matrix of S under Q :

$$C_t^{ij} = \mathbb{E}^Q[S^i S^j | \mathcal{F}_t] - \mathbb{E}^Q[S^i | \mathcal{F}_t] \mathbb{E}^Q[S^j | \mathcal{F}_t].$$

We may have these simple expressions for S^* and C^{ij} because the optimal strategy is of buy-and-hold type.

The boundedness condition on S , G and H can be relaxed. Due to (16), standard models of mathematical finance are then supported in this framework. For example, if $d = 1$ and

$$S = \mu + \sigma W_1, \quad H = \alpha S, \quad G = \beta S$$

with a standard Brownian motion W and constants $\alpha, \beta, \mu \in \mathbb{R}$ and $\sigma > 0$, then we obtain the Bachelier model

$$dS_t^* = \frac{c\gamma}{c + \gamma}(\alpha + \beta)\sigma^2 dt + \sigma dW_t, \quad S_0^* = \mu - \frac{c}{c + \gamma}(\alpha + \beta)\sigma^2.$$

If $d = 1$ and

$$S = \zeta e^{\sigma W_1}, \quad G = \alpha \sigma W_1 - H, \quad H = \frac{\gamma}{c + \gamma} \alpha \sigma W_1 - \frac{\mu}{c + \gamma} S$$

with constants $\alpha, \mu \in \mathbb{R}$ and $\sigma, \zeta > 0$, then we have $Y_t^* = \mu/(c + \gamma)$ and obtain the Black-Scholes model

$$dS_t^* = \frac{c\gamma}{c + \gamma} \alpha \sigma^2 S_t^* dt + \sigma S_t^* dW_t, \quad S_0^* = \zeta \exp \left\{ \sigma^2 \left(\frac{1}{2} - \frac{c\gamma}{c + \gamma} \alpha \right) \right\}.$$

5.3 Complete Markov models revisited

Here we reconsider the framework of Section 4.4 in terms of the dynamic programming principle. Our aim here is to recover Theorem 2 by letting $n \rightarrow \infty$. The filtration $\{\mathcal{F}_t\}$ is generated by a k -dimensional $\{\mathcal{F}_t\}$ -standard Brownian motion W . We suppose that A is a compact set and

$$S = s(W_1), \quad G = g(W_1), \quad H = h(W_1)$$

with functions $s : \mathbb{R}^k \rightarrow \mathbb{R}^d$, $g : \mathbb{R}^k \rightarrow \mathbb{R}$ and $h : \mathbb{R}^k \rightarrow \mathbb{R}$. Here we assume in addition the boundedness and the uniform continuity for the derivatives

$$\frac{\partial v}{\partial t}, \quad \frac{\partial v}{\partial w_i}, \quad \frac{\partial^2 v}{\partial w_i \partial w_j}, \quad \frac{\partial p}{\partial t}, \quad \frac{\partial p}{\partial w_i}, \quad \frac{\partial^2 p}{\partial w_i \partial w_j}, \quad i, j = 1, \dots, k.$$

for p and v defined in Section 4.4.

Lemma 6 *Under the completeness condition (14),*

$$\lim_{n \rightarrow \infty} n \max_{t \in D_n} \|v(t, W_t) - \Psi_t(v(t + 1/n), W_{t+1/n})\|_\infty = 0.$$

Proof: By (13),

$$p(t, W_t, y) = \Pi_t(p(t + 1/n, W_{t+1/n}, y)).$$

By Taylor's theorem,

$$\begin{aligned} & p(t, w, y) \\ &= -\frac{1}{\gamma} \log \mathbb{E} \left[\exp(-\gamma p(t + 1/n, W_{t+1/n}, y)) \mid W_t = w \right] \\ &= p(t + 1/n, w, y) \\ &\quad - \frac{1}{\gamma} \log \left\{ 1 - \frac{\gamma}{2n} \Delta p(t + 1/n, w, y) + \frac{\gamma^2}{2n} |\nabla p(t + 1/n, w, y)|^2 + o(1/n) \right\} \\ &= p(t + 1/n, w, y) + \frac{1}{2n} \Delta p(t + 1/n, w, y) - \frac{\gamma}{2n} |\nabla p(t + 1/n, w, y)|^2 + o(1/n). \end{aligned}$$

Here and hereafter, the estimate $o(1/n)$ is uniform in (t, w, y) . Similarly,

$$\begin{aligned} & -\frac{1}{c} \log \mathbb{E} \left[\exp(-c(v(t + 1/n, W_{t+1/n}) - p(t + 1/n, W_{t+1/n}, y))) \mid W_t = w \right] \\ &= v(t + 1/n, w) - p(t + 1/n, w, y) + \frac{1}{2n} (\Delta v(t + 1/n, w) - \Delta p(t + 1/n, w, y)) \\ &\quad - \frac{c}{2n} |\nabla v(t + 1/n, w) - \nabla p(t + 1/n, w, y)|^2 + o(1/n). \end{aligned}$$

Therefore,

$$\begin{aligned}
& \Psi_t(v(t+1/n, W_{t+1/n})) \\
&= \sup_{y \in A} \left\{ v(t+1/n, W_t) + \frac{1}{2n} \Delta v(t+1/n, W_t) - \frac{c}{2n} |\nabla v(t+1/n, W_t)|^2 \right. \\
&\quad + \frac{c}{n} (\nabla v(t+1/n, W_t), \nabla p(t+1/n, W_t, y)) \\
&\quad \left. - \frac{c+\gamma}{2n} |\nabla p(t+1/n, W_t, y)|^2 + o(1/n) \right\} \\
&= v(t, W_t) - \inf_{y \in A} \left\{ \frac{c+\gamma}{2n} \left| \nabla p(t+1/n, W_t, y) - \frac{c}{c+\gamma} \nabla v(t+1/n, W_t) \right|^2 + o(1/n) \right\}.
\end{aligned}$$

Then the result follows from the completeness condition. ////

Theorem 3 *Under the completeness condition (14),*

$$\lim_{n \rightarrow \infty} V_0^n(0, 0) = v(0, W_0) - p(0, W_0, 0) = U_0^*(G + H) - \Pi_0(G).$$

Proof: Put

$$r_n = \sup_{t \in D_n} \|v(t, W_t) - \Psi_t(v(t+1/n), W_{t+1/n})\|_\infty.$$

By the previous lemma, $nr_n \rightarrow 0$ as $n \rightarrow \infty$. For $k = 0, 1, \dots, n-1$, put

$$F_k = \Psi_{k/n} \circ \Psi_{(k+1)/n} \circ \dots \circ \Psi_{(n-1)/n}(G + H).$$

Then, by the cash-invariance property of $\Psi_{k/n}$,

$$\begin{aligned}
\|F_k - v(k/n, W_{k/n})\|_\infty &\leq \|\Psi_{k/n}(F_{k+1}) - \Psi_{k/n}(v((k+1)/n, W_{(k+1)/n}))\|_\infty \\
&\quad + \|\Psi_{k/n}(v((k+1)/n, W_{(k+1)/n})) - v(k/n, W_{k/n})\|_\infty \\
&\leq \|F_{k+1} - v((k+1)/n, W_{(k+1)/n})\|_\infty + r_n.
\end{aligned}$$

Then, by Proposition 2,

$$\begin{aligned}
& |V^n(0, 0) - v(0, X_0) + p(0, X_0, 0)| \\
&= |F_0 - v(0, X_0^n)| \leq nr_n \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$. ////

References

- [1] Y. Amihud and H. Mendelson. Dealership market. *Journal of Financial Economics* 8, 31-53, 1980.
- [2] P. Bank and D. Kramkov. A model for a large investor trading at market indifference prices. I: single-period case. arXiv:1110.3224v3, 2013.

- [3] P. Bank and D. Kramkov. A model for a large investor trading at market indifference prices. II: continuous-time case. arXiv:1110.3229v3, 2014.
- [4] P. Barrieu and N. El Karoui. Inf-convolution of risk measures and optimal risk transfer. *Finance Stoch.* 9, no. 2, 269-298, 2005.
- [5] U. Cetin, R. A. Jarrow, and P. Protter. Liquidity risk and arbitrage pricing theory. *Finance Stoch.* 8(3):311341, 2004.
- [6] F. Delbaen. *Monetary utility functions*. Osaka University CSFI Lecture Notes Series 3, Osaka University Press, 2012.
- [7] M. Fukasawa. Efficient discretization of stochastic integrals. *Finance Stoch.* 18, 175-208. 2014.
- [8] M.B. Garman. Market microstructure. *Journal of Financial Economics* 3, 257-275, 1976.
- [9] L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14(1):71 100, 1985.
- [10] O. Guéant. Permanent market impact can be nonlinear. arXiv:1305.0413v4, 2014.
- [11] T. Ho and H. R. Stall. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9, 47-73, 1981.
- [12] Hopf, Eberhard. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Comm. Pure Appl. Math.* 3, (1950). 201-230.
- [13] J. Kallsen and A. Shiryaev. The cumulant process and Esscher's change of measure. *Finance Stoch.* 6, no. 4, 397-428, 2002.
- [14] M. O'hara and G.S. Oldfield. The Microeconomics of Market Making. *The Journal of Financial and Quantitative Analysis* 21 (4), 361-376, 1986.
- [15] K.R. Parthasarathy. *Probability measures on metric spaces*. Probability and Mathematical Statistics, No. 3 Academic Press, Inc., New York-London, 1967.
- [16] P.E. Protter. *Stochastic integration and differential equations*. Second edition. Stochastic Modelling and Applied Probability, 21. Springer-Verlag, Berlin, 2005.
- [17] K. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge Studies in Advanced Mathematics, 68. Cambridge University Press, Cambridge, 1999.